WORKING PAPERS

W.P. n. 21

CALIBRATING THE RESIDENTIAL LOCATION SUBMODEL OF THE SIMULATION MODEL FOR THE TURIN METROPOLITAN AREA

C.S. Bertuglia (*), I. Gallino (*), I. Gualco (*) S. Occelli (*), G.A. Rabino (*), C. Salomone (**), R. Tadei (*)

ires ISTITUTO RICERCHE ECONOMICO - SOCIALI DEL PIEMONTE

[redacted]

W.P. n. 21

CALIBRATING THE RESIDENTIAL LOCATION SUBMODEL OF THE SIMULATION MODEL FOR THE TURIN METROPOLITAN AREA

C.S. Bertuglia (*), I. Gallino (*), I. Gualco (*)
S. Occelli (*), G.A. Rabino (*), C. Salomone (**),
R. Tadei (*)

Giugno 1983

(*) IRES – Istituto di Ricerche Economico Sociali del Piemonte, Torino, Via Bogino n. 21.

(**) CERIS – Istituto di Ricerca sull'Impresa e lo Sviluppo del Consiglio Nazionale delle Ricerche, Torino, Via Avogadro n. 8.

Paper prepared for the AIRO Conference 1983, Naples, September 26-28.

SUMMARY

In this paper we discuss the main calibration results of the residential location submodel which is part of a comprehensive simulation model, now being applied to the Turin metropolitan area. The theoretical structure and the procedure for the calibration of the simulation model were described at the Airo Conference in 1982.

Building upon these results we propose both the methodological and operational developments for this residential location submodel.

Further results obtained from testing some of the suggested developments are also presented.

Contents

1. Introduction

In this paper we discuss the main calibration results of the residential location submodel which is part of a comprehensive simulation model now being applied to the Turin metropolitan area.
The theoretical structure and the calibration procedure of the simulation model, as well as some results of simulation experiments were presented at the Airo Conference in 1982 (Bertuglia, Gallino, Gualco, Occelli, Rabino, Salomone, Tadei, 1982a).

Here, we mainly focus on the residential location submodel, as at this stage, it is the one which creates the most difficult but most stimulating calibration problems both from the theoretical and operational points of view.
In the following sections the methodological and operational directions for development of the submodel are put forward and the results of the investigation of some of these alternatives are presented.

2. Characteristics and problems in calibrating the residential location submodel

2.1. Structure of the residential location submodel

In order to clarify the following discussion, we restate the analytical structure of the residential location submodel. To simplify the notation we omit the indices relative to zones (i,j = 1,99), family type (f = 1,8), housing type (s = 1,6), industrial and service sections (t = 1,4, l = 1) and transport mode (v = 1,2). Let us define

O   total number of jobs (calculated in the industry and service submodels);

KO inverse of the mean rate of employment of households;

Q = O/KO total number of households (with employed family heads);

TPROB probability of utilization of transport mode;

T travel costs (transportation submodel);

AB housing (calculated in the housing submodel);

SLAO land in residential use (calculated in the land use submodel);

K, H, N weighting factors;

TETA, CSI parameters;

$$A = \Sigma Q \cdot TPROB \cdot e^{-TETA \cdot T} \quad \text{residential accessibility.} \qquad (1)$$

The utility (real utility) derived by a household from the choice of a residential bundle (residential zone and housing type) takes the form:

$$U = K \cdot \overline{A} + H.\overline{AB} + N \cdot \overline{SLAO} \qquad (2)$$

(where $\overline{x}$ is the normalised value of x).

The expected value of utility $\overline{U}$ is given by:

$$\overline{U} = \Sigma\, U\, \Sigma\, \frac{DPOTO}{Q} \qquad (3)$$

where

$$DPOTO = Q \cdot \frac{WT}{\Sigma WT} \qquad (4)$$

and

$$WT = \Sigma TPROB \cdot e^{-TETA \cdot T} \cdot \Sigma\, e^{-CSI\,(\overline{U}-U)}. \qquad (5)$$

DPOTO is the representative variable of both the residential and workplace location of families (broken down by family type).

Model (1) (5) is based on the hypothesis that not all the families are likely to find an optimal residential location: some attain a higher utility level than expected, while others remain below it.

In this way, the model tries to describe a real market, providing a measure of demand and supply disequilibrium based on the difference between the real and the expected utility of families.

A diagrammatic representation of model (1) - (5) is given in fig. 1.

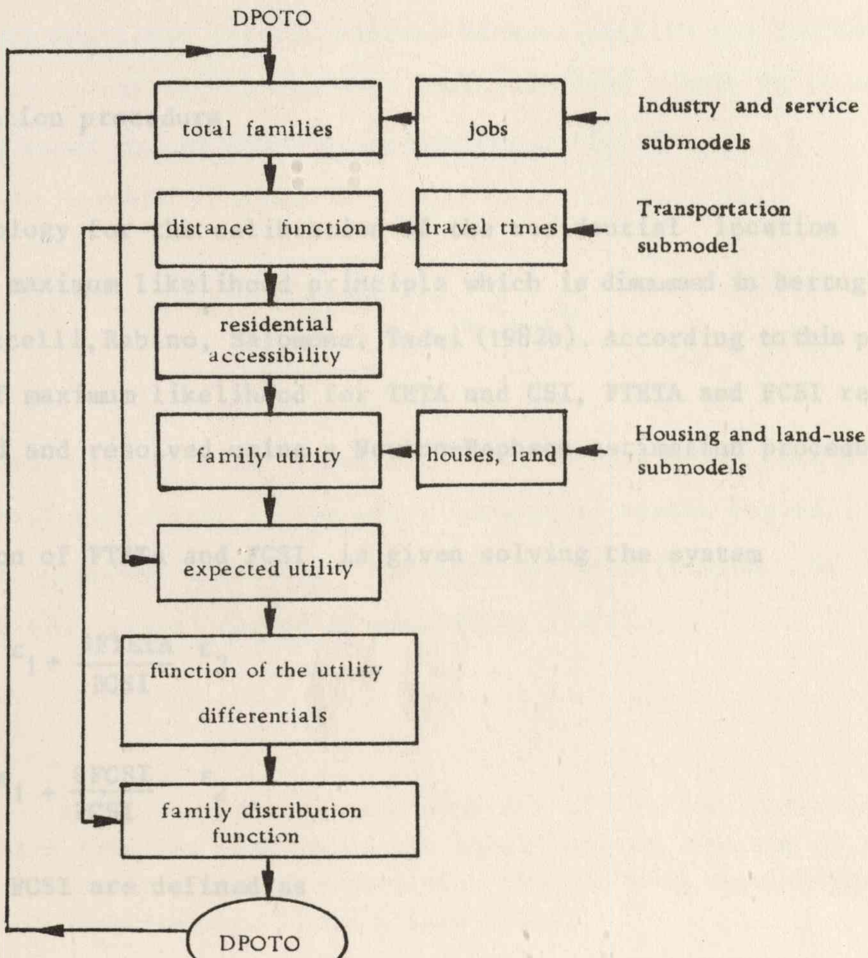

Figure 1 - Structure of the residential location submodel

It can be seen that the residential distribution of families depends on two main factors: a spatial factor (distance function) and a factor measuring family welfare (function of the utility differentials).

The relative importance of these factors and thus the level of competitivity which is likely to exist between them is expressed by the parameters TETA and CSI.

In addition, family utility depends, among other factors, on the residential accessibility [from eq.(1)], the exponential term of which appears also in the distribution function of DPOTO [eqs. (4) - (5)].

2.2. The calibration procedure

The methodology for the calibration of the residential location submodel was based on the maximum likelihood principle which is discussed in Bertuglia, Gallino, Gualco, Occelli, Rabino, Salomone, Tadei (1982b). According to this principle two equations of maximum likelihood for TETA and CSI, FTETA and FCSI respectively, were derived and resolved using a Newton-Raphson estimation procedure.

The solution of FTETA and FCSI is given solving the system

$$-\mathrm{FTETA} = \frac{\partial \mathrm{FTETA}}{\partial \mathrm{TETA}} \varepsilon_1 + \frac{\partial \mathrm{FTETA}}{\partial \mathrm{CSI}} \varepsilon_2 \tag{6}$$

$$-\mathrm{FCSI} = \frac{\partial \mathrm{FCSI}}{\partial \mathrm{TETA}} \varepsilon_1 + \frac{\partial \mathrm{FCSI}}{\partial \mathrm{CSI}} \varepsilon_2 , \tag{7}$$

where FTETA and FCSI are defined as

$$FTETA = C - C^{obs} \tag{8}$$

$$FCSI = \Delta U - \Delta U^{obs} \tag{9}$$

and ε_1, ε_2 are the differences between the best and approximated values of the parameters;

C, C^{obs} are, respectively, the predicted and observed mean cost functions;

$\Delta U, \Delta U^{obs}$ are, respectively, the predicted and observed functions of the utility differentials.

The main feature which characterizes the calibration procedure is that, because of the above mentioned interdependence between utility and residential accessibility, it was necessary to nest the Newton-Raphson scheme in an outer iterative structure based on the Hyman procedure (cf.: fig. 2).

The converging formula of Hyman which has been used here is:

$$TETA_{m+1} = TETA_m + \left[(C^{obs} - C_m)(TETA_m - TETA_{m-1}) / (C_m - C_{m-1}) \right] . \tag{10}$$

As far as the data for this submodel calibration are concerned, it was assumed that, without variations in the other submodels (static regime), the model (1) - (5) should reproduce accurately the observed distribution of families (DPOTO) (*), in the initial period of simulation (1971).

(*) The observed DPOTO in the base year consists of 470,488 possible combinations resulting from the product of the model indices. The sum of the values of all the combinations comes to just over 500,000 thus, showing that many of the possible combinations have a zero value.

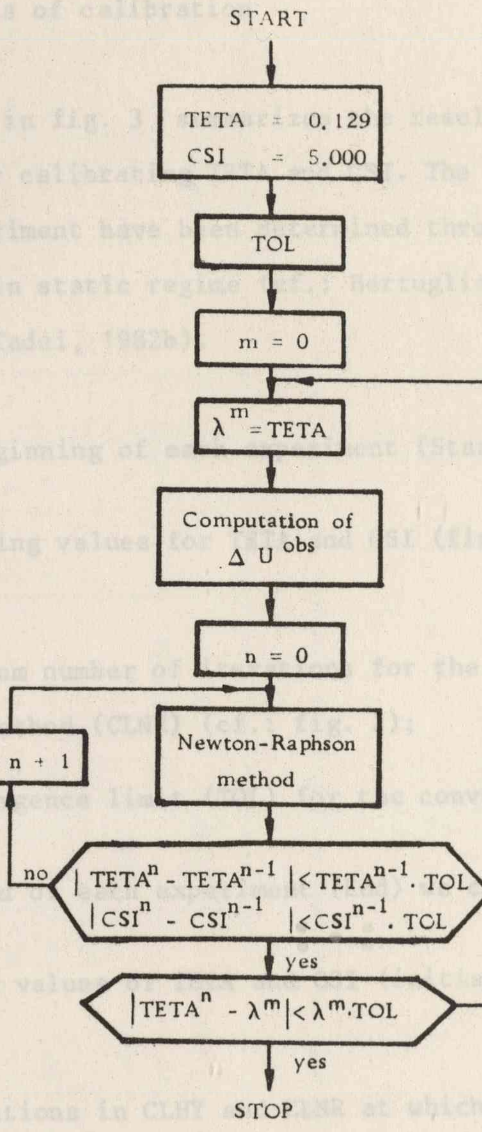

Figure 2 - Diagram of the calibration procedure

2.3. The results of calibration

The table in fig. 3 summarizes the results of the experiments which were carried out for calibrating TETA and CSI. The initial values of TETA and CSI in the first experiment have been determined through a sensitivity analysis of the general model in static regime (cf.: Bertuglia, Gallino, Gualco, Occelli, Rabino, Salomone, Tadei, 1982b).

At the beginning of each experiment (Start) we define:

a. the starting values for TETA and CSI (final values of the previous experiment);

b. the maximum number of iterations for the Hyman procedure (CLHY) and Newton-Raphson method (CLNR) (cf.: fig. 2);

c. the convergence limit (TOL) for the convergence texts in CLHY and CLNR.

At the end of each experiment (End) we calculate:

d. the final values of TETA and CSI (initial values of the following experiments);

e. the iterations in CLHY and CLNR at which the convergence is attained;

f. the values of the maximum likelihood equations for TETA (FTETA), eq.(6), and CSI (FCSI), eq. (7);

g. an indicator of the deviation between the predicted DPOTO (CAL) and observed DPOTO (OB) which is expressed by:

$$SCARTO = [(CAL-OB)^2/N]^{\frac{1}{2}} \qquad (11)$$

where N (470,488) is the product of the model indices.

	1		2		3		4		5		6		7		8	
	Start	End	Start	End	Start	End	Start	End	Start	End	Start	End	Start	End	Start	End
TETA	0.129	0.1054 →		0.1042 →		0.10367 →		0.10361 →		0.103607 →		0.1036069 →		0.10360687 →		0.10360685
CSI	5.	5.731 →		5.7421 →		5.7781 →		5.7788 →		5.7793 →		5.779329 →		5.77933 →		5.779334
CJHY	10	10	10	2	10	10	30	4	30	11	30	2	30	11	30	1
CLNR	20	20	20	20	30	30	30	18	30	30	30	30	30	30	30	10
TOL	10^{-5}		10^{-5}		10^{-6}		10^{-6}		10^{-8}		10^{-11}		10^{-11}		10^{-11}	
FTETA		-243121		-75752		-8503		-1373		-39		-6		-1.43		-0.94
FCSI		699		488		19		6		0.1124		0.6595		0.1614		0.2382
SCARTO		7.100		7.122		7.1433		7.1446		7.14471		7.14477		7.14478		7.14478

Figure 3 - Table of the experiments performed for the calibration of TETA an CSI

TETA, CSI — values of the parameters

CIHY, CLNR — iterations for the Hyman procedure and the Newton-Raphson method

TOL — convergence limit

FTETA, FCSI — values of the maximum likelihood equations for TETA and CSI

SCARTO — deviation between the predicted (CAL) and observed DPOTO (OB)

Different experiments have been run within a sequence of execution based on the following.

If an experiment with given starting definitions, does not reach the convergence the experiment is reiterated updating TETA and CSI (and eventually changing CLHY and CLNR).

If the convergence is attained, but the values of FTETA and FCSI are not sufficiently small a new experiment is defined.

The best values of TETA and CSI, were found after eight experiments. Each experiment consisted of the number of runs given by product of the values of CLHY and CLNR listed in fig. 3 (End).

As shown in fig. 3, the adjustment process of the TETA and CSI is performed mainly during the first five experiments, while remaining substantially unchanged in the following ones. However, only in the eighth experiment, did both the FTETA and FCSI approximate the zero value.

Moreover it should be noted that the difference between the predicted and observed DPOTO as defined in eq. (11), increases in all the experiments (starting, however, from a high initial value).

Fundamentally these results show that although from a strictly methodological point of view (mimisation of FTETA and FCSI), the goodness of fit can be considered satisfactory, the increase of the above difference seems to a certain extent to bias its level of accuracy.

For the best values of the parameters a further analysis of the predicted

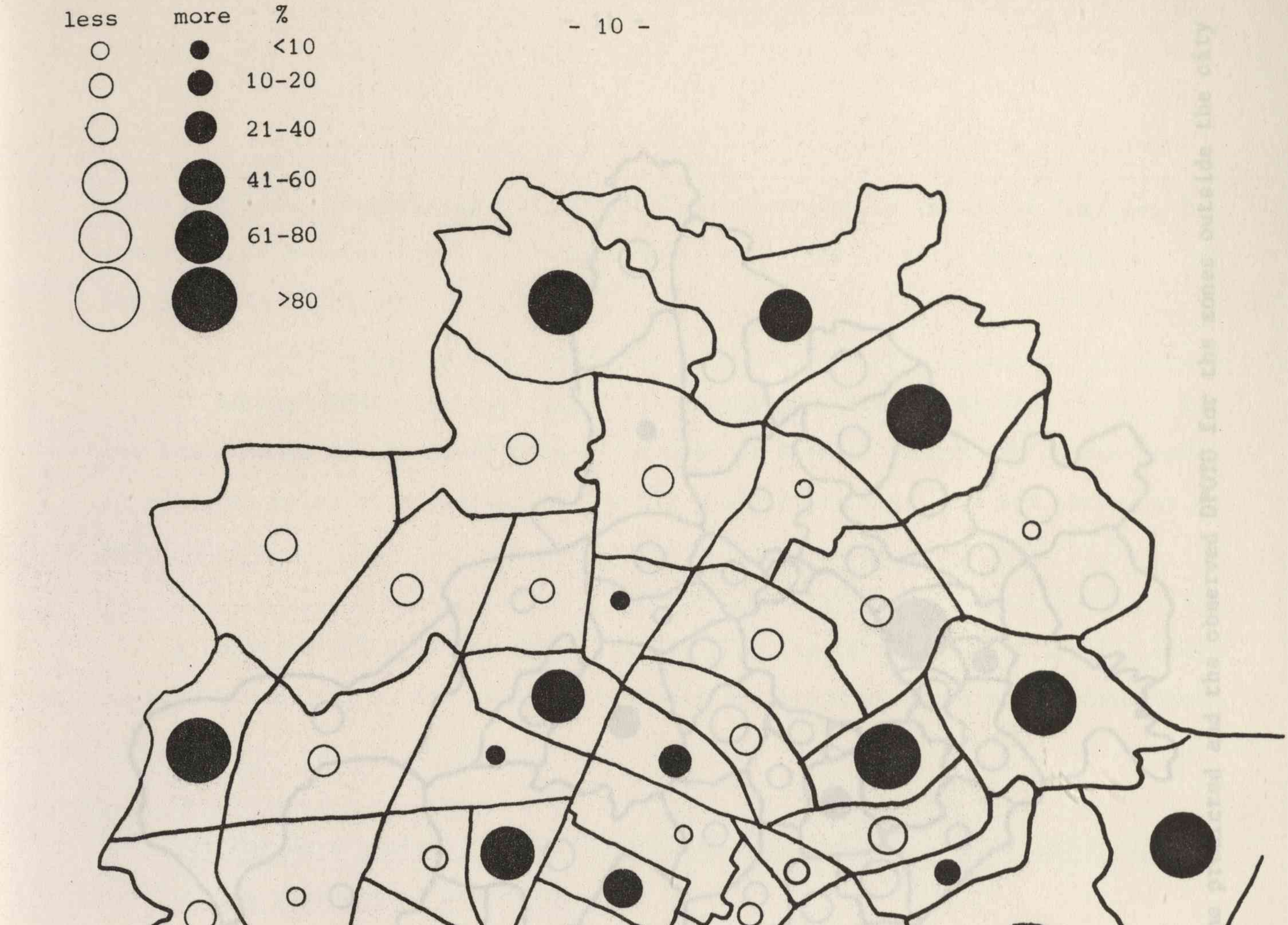

Figure 4 - Percentage difference between the predicted and the observed DPOTO for the zones in the city of Turin

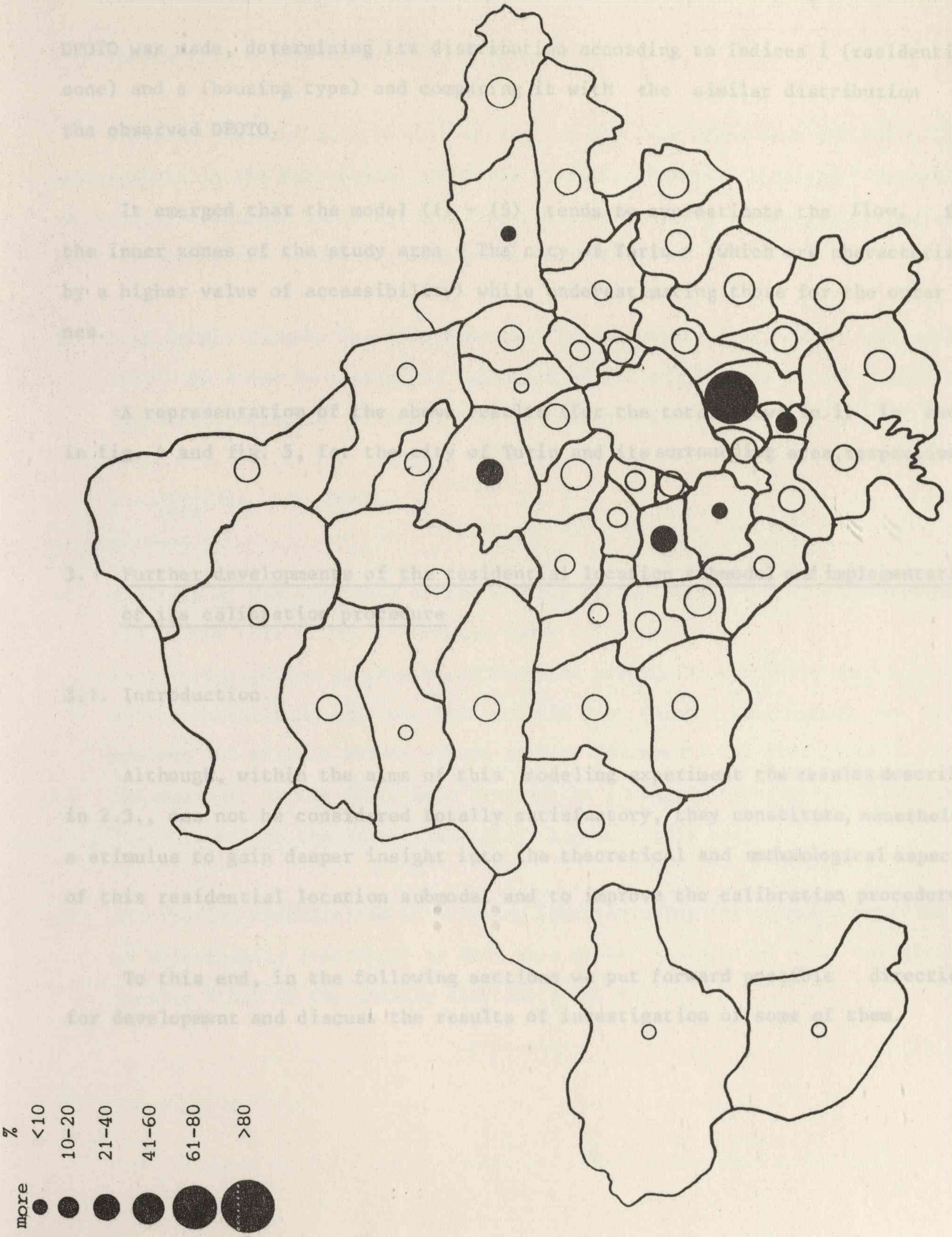

Figure 5 - Percentage difference between the predicted and the observed DPOTO for the zones outside the city of Turin

DPOTO was made, determining its distribution according to indices i (residential zone) and s (housing type) and comparing it with the similar distribution of the observed DPOTO.

It emerged that the model (1) - (5) tends to overestimate the flow, for the inner zones of the study area - The city of Turin - (which are characterized by a higher value of accessibility) while underestimating those for the outer zones.

A representation of the above results (for the total flows in i) is shown in fig. 4 and fig. 5, for the city of Turin and its surrounding area, respectively.

3. Further developments of the residential location submodel and implementation of its calibration procedure

3.1. Introduction

Although, within the aims of this modeling experiment the results described in 2.3., can not be considered totally satisfactory, they constitute, nonetheless, a stimulus to gain deeper insight into the theoretical and methodological aspects of this residential location submodel and to improve the calibration procedure.

To this end, in the following sections we put forward possible directions for development and discuss the results of investigation of some of them.

3.2. Theoretical and methodological developments

Building upon the main results of section 2.3., we argue that the following developments in the theoretical structure of the residential location submodel can provide a sounder basis for modeling the residential location process, not least in improving the accuracy of the goodness-of-fit.

a. Logarithmic formulation of residential accessibility [eq. (1)] (Leonardi, 1979) (in order to lessen the impedance effect of distance in the utility function). Equation (1) could then be restated as

$$A = \frac{1}{TETA} \log (\Sigma\, Q \cdot TPROB \cdot e^{-TETA \cdot T}) . \qquad (12)$$

b. Analysis of the structure of the utility differentials ($\bar{U}$-U), and consequently of the role of CSI in varying these differentials.

Some investigations in this direction have already been carried out, introducing euristically (for the TETA and CSI best value), some bounds on the maximum and minimum values of the utility differentials. This allowed for the decrease in the difference between the predicted and observed DPOTO from 7.144 to 6.390, the latter remaining, however still high.

c. Alternative formulations of the utility function (using for example, logarithmic or Cobb-Douglas functions) as well as a deeper analysis of the weighting factors K,H,N in the utility function [eq. (2)].

3.3. Implementation of the calibration procedure

At this stage of the submodel application, the following modifications and integrations in the calibration operations can be useful to improve both the estimates of the predicted DPOTO and the overall efficiency of the procedure itself.

a. Elimination of the Hyman procedure (outer loop in fig. 2), using only the Newton-Raphson sequence for the simultaneous computation of TETA and CSI as well as for the updating of TETA in the residential accessibility function.

b. Disaggregation of the TETA and CSI parameters by family types (that is eight TETA's and eight CSI's), as in the original version of the submodel (Bertuglia, Occelli, Rabino, Tadei, 1980).

c. Introduction in the DPOTO distribution function [eq. (5)] of corrective factors associated with destinations (residential zone and housing type). Some first experiments have been performed, and these factors (with starting unit values) have been calculated through a trial and error iterative procedure.

Taking into account these factors, the recomputation of the difference [eq. (11)] showed a small decrease from 6.390 to 6.068.

In fact, these factors can be given a more sound theoretical interpretation if they are considered:

c.a. as a measure of the attractiveness associated with each destination i,s. In this case they should be explicitly introduced in eq. (5) and determined through an appropriate calibration procedure;

c.b. as weighting factors associated with the utility differentials thus playing a twofold role (Anas, 1973):

1. modifying the entropy distribution which otherwise would be continuous over all the values of the utility function;

2. rectifying some inconsistencies which might arise in the estimated values of the utility differentials.

3.4. Alternative directions of development

We outline in fig. 6 the general framework of the alternative directions which could be investigated for the future development of the residential location submodel, as discussed in 3.2. and 3.3..

theoretical developments \ operational implementations	a unique Newton Raphson iteration	b 8 TETA's and 8 CSI's	c factors associated to destinations
a. no developments	X_1	X_2	-
b. logarithmic accessibility function	X_3	X_4	X_5
c. analysis of the utility differentials	X_6	X_7	X_8
d. alternative formulations of the utility functions	X_9	X_{10}	X_{11}

Figure 6 - Framework of the alternative directions of development /where X_i (i=1,11) represents the alternative which could be investigated /

3.5. Outcome of tested alternatives

So far, alternative X_1 has been explored (cf.:fig.6). The procedure was started from the best values of the parameters TETA and CSI we found with the Hyman/ Newton-Raphson iterations (cf.: fig. 3). (Each iteration took 6.2 second of CPU time).

However, it quickly diverged to extreme values thus implying that these parameter values are not the best starting values.

A systematic exploration of these starting values is then necessary.

Of course, the unique Newton-Raphson sequence might turn out to be less satisfactory than the combined Hyman/Newton-Raphson procedure. This is a point which should be assessed and also requires the investigation of the theoretical alternatives previously suggested, cf.: fig. 6.

4. Conclusions

This paper dealt with the theoretical and operational problems of calibrating the residential location submodel which is part of a comprehensive simulation model now being applied to the Turin metropolitan area.

Because of the interdependence between the utility function and the residential accessibility the simultaneous calibration of the submodel parameters required a nested iterative procedure the outer loop of which- the Hyman procedure - embedded a Newton-Raphson scheme.

Building upon the results of the calibration operations the theoretical and operational developments for this submodel were highlighted.

Finally, in discussing the first outcomes of the tested alternative there emerged a feeling that improvements of this submodel calibration will need an accurate matching between the theoretical and operational developments which could be carried out

References

Anas A. (1973) A Dynamic Disequilibrium Model of Residential Location, Environment and Planning A, 5, 633-647.

Bertuglia C.S., Gallino T., Gualco I., Occelli S., Rabino G.A., Salomone C., Tadei R. (1982a) Alcuni aspetti della calibrazione di un modello dinamico spazializzato: il caso del modello dell'area metropolitana torinese, Atti delle Giornate di lavoro AIRO 1982, 200-248.

Bertuglia C.S., Gallino T., Gualco I., Occelli S., Rabino G.A., Salomone C., Tadei R. (1982b) L'applicazione di un modello dinamico a larga scala per l'Area Metropolitana di Torino: la calibrazione, WP 14, IRES, Torino (comunicazione presentata alla III^ Conferenza Nazionale di Scienze Regionali, Venezia, 10-12/12/1982).

Bertuglia C.S., Occelli S., Rabino G.A., Tadei R. (1980) A Model of Urban Structure and Development of Turin: Theoretical Aspects, Sistemi Urbani, 2, 59-90.

Leonardi G. (1979) Introduzione alla teoria dell'accessibilità, Sistemi Urbani, 1 1, 65-87.

WORKING PAPERS

*1 "Un modello urbano a larga scala per l'area metropolitana di Torino", *gennaio 1981*

*2 "Metodologie per la pianificazione dei parchi regionali", *gennaio 1981*

*3 "A Large Scale Model for Turin Metropolitan Area", *maggio 1981*

4 "An Application to the Ticino Valley Park of a Mathematical Model to Analyse the Visitors Behaviour", *luglio 1981*

5 "Applicazione al parco naturale della Valle del Ticino di un modello per l'analisi del comportamento degli utenti: la calibrazione del modello", *settembre 1981*

6 "Applicazione al parco naturale della Valle del Ticino di un modello per l'analisi del comportamento degli utenti: l'uso del modello", *settembre 1981*

*7 "Un'analisi delle relazioni esistenti tra superficie agricola utilizzata ed alcune principali grandezze economiche in un gruppo di aziende agricole piemontesi al 1963 e al 1979", *settembre 1981*

8 "Localizzazione ottimale dei servizi pubblici, con esperimenti sulle scuole dell'area torinese", *settembre 1981*

9 "La calibrazione di un modello a larga scala per l'area metropolitana di Torino", *ottobre 1981*

10 "Applicazione al parco naturale della Valle del Ticino di un modello per l'analisi del comportamento degli utenti: l'individuazione di un indicatore di beneficio per gli utenti ed una analisi di sensitività su alcuni parametri fondamentali", *ottobre 1981*

11 "La pianificazione dell'uso ricreativo di aree naturali: il caso del parco della Valle del Ticino", *novembre 1981*

*12 "The Recreational Planning of Country Parks: the Case Study of the Ticino Valley Park", *marzo 1982*

13 "Alcuni aspetti della calibrazione di un modello dinamico spazializzato: il caso del modello dell'area metropolitana torinese", *settembre 1982*

*14 "L'applicazione di un modello dinamico a larga scala per l'area metropolitana di Torino: la calibrazione", *novembre 1982*

15 "Modello commerciale Piemonte", *novembre 1982*

16 "Resource allocation in multi-level spatial health care systems: benefit maximisation", *dicembre 1982*

17 "Relazione sulla struttura e sulla dinamica del settore elettromeccanico piemontese", *dicembre 1982*

18 "Evoluzione della finanza locale in Piemonte e in Italia 1977 - 1981", *febbraio 1983*

19 "Un metodo per l'analisi di scenari multidimensionali in ordine alle relazioni tra domanda di trasporto e variabili strutturali dei sistemi economici e territoriali", *febbraio 1983*

20 "Modello commerciale Piemonte", *marzo 1983*

---

* Esaurito

ires
ISTITUTO RICERCHE ECONOMICO - SOCIALI DEL PIEMONTE
VIA BOGINO 21 10123 TORINO